Research Article

TRANSPARENT PROCESS OPEN ACCESS

Life Science Alliance

pSATdb: a database of mitochondrial common, polymorphic, and unique microsatellites

Sonu Kumar[1], Ashutosh Singh[2], Asheesh Shanker[1]

Microsatellites, also termed as simple sequence repeats, are repetitive tracts in a DNA sequence, typically consisting of one to six nucleotides. These repeats are found in all genomes and play key roles in phylogeny and species identification. Microsatellites are highly polymorphic, and their length may differ from species to species. There are several online resources dedicated to mitochondria; however, comprehensive information is not available about the length variation of mitochondrial microsatellites. Therefore, to explore it between species among a genus, we have developed a database named pSATdb (polymorphic microSATellites database; https://lms.snu.edu.in/pSATdb/). pSATdb contains 28,710 perfect microsatellites identified across 5,976 mitochondrial genome (mt-genome) sequences from 1,576 genera which includes 1,535 (5,846 mt-genome) and 41 (130 mt-genome) genera of Metazoa and Viridiplantae, respectively. pSATdb is the only database which provides genus-wise information about the length variation of mitochondrial microsatellites. Because of the emerging role of microsatellites in genomics studies, the identified common, polymorphic, and unique microsatellites stored in pSATdb will be effectively useful in various studies including genetic diversity, mapping, marker-assisted selection, and comparative population studies.

DOI 10.26508/lsa.202101307 | Received 19 November 2021 | Revised 3 February 2022 | Accepted 4 February 2022 | Published online 18 February 2022

Introduction

Mitochondria, often referred to as the "powerhouses of the cell," are an essential cellular organelle found in eukaryotes. Mitochondria possess its own genome and plays essential roles in cellular respiration (Roger et al, 2017), phylogeny (Kern et al, 2020), and species identification (Yang et al, 2014; Stoeckle & Thaler, 2018). Mitochondrial genome (mt-genome) contains repetitive sequences including microsatellites with varying lengths (Habano et al, 1998).

Microsatellites, also termed, as simple sequence repeats (SSRs), are a repetitive tract in DNA, typically consisting of one to six nucleotides (Tautz & Renz, 1984). Based on the composition of the repeats, microsatellites were categorized as perfect, imperfect, and compound microsatellites. Repeats without interruption are known as perfect microsatellites (e.g., AAAAAAAA), whereas imperfect microsatellites are interrupted by non-repeat nucleotides (e.g., AAAA***T***AAAA). Two or more microsatellites found adjacent to each other or separated by few nucleotides are called compound microsatellites (e.g., AAAAAAAA*TTTTTTTT*; Bachmann et al, 2004). These repeats are found in coding, non-coding, and coding–non-coding regions of both eukaryotic and prokaryotic genomes (Shanker et al, 2007a, 2007b; Kapil et al, 2014; Kabra et al, 2016; Kumar & Shanker, 2018a). Moreover, these repeats have also been reported in organellar genome including mitochondria (Kumar et al, 2014, 2020; Kumar & Shanker, 2020a).

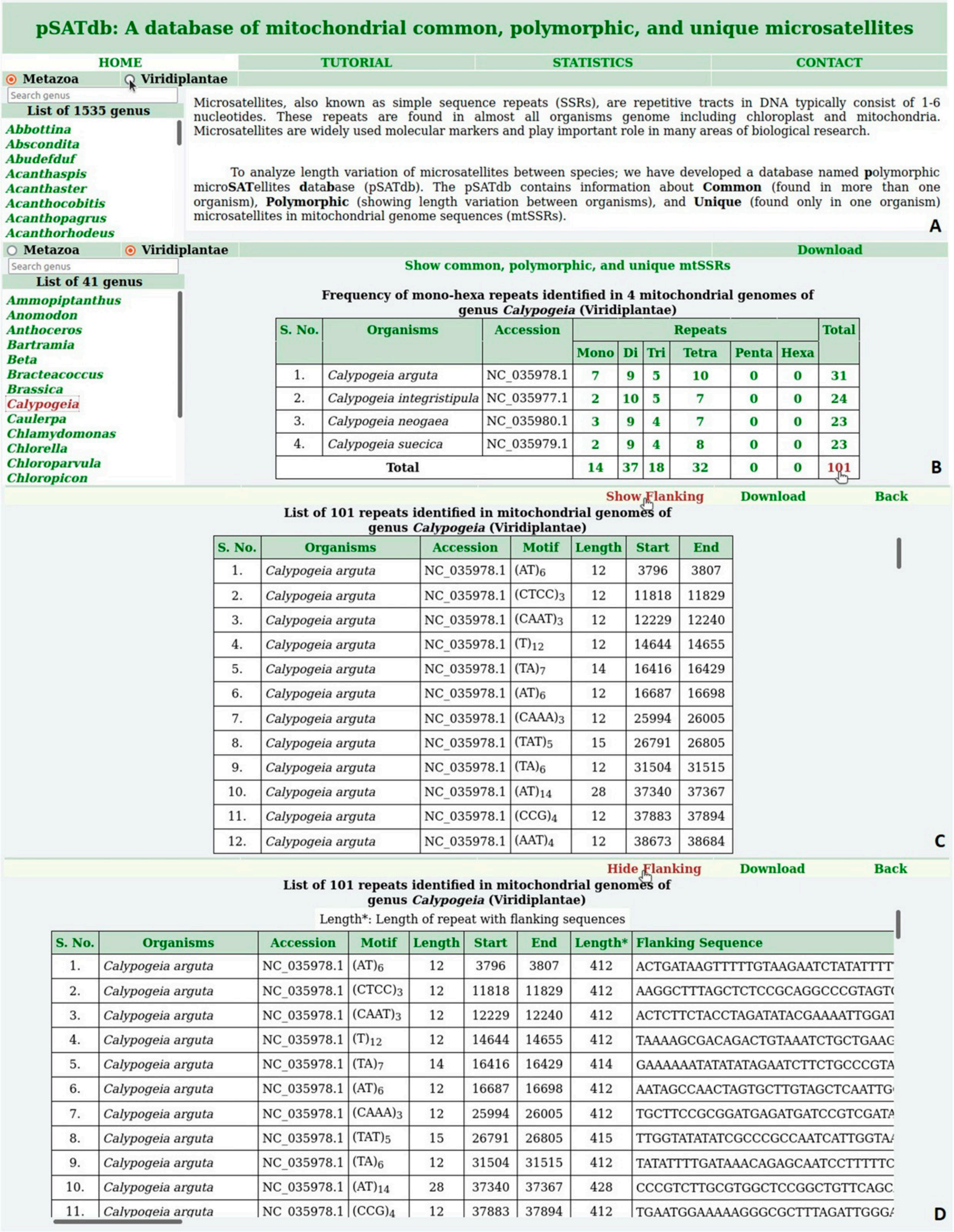
Microsatellites have been widely applied as a powerful genetic/molecular marker because of their abundance, high reproducibility, hypervariability, codominant and multi-allelic nature (Powell et al, 1996; Parida et al, 2009). Consequently, microsatellites were applied for a variety of purposes including genetic, evolutionary, molecular breeding, and phylogenetic studies (Agarwal et al, 2008; Stolle et al, 2013; Deng et al, 2016; Fu et al, 2016). Earlier, studies were conducted to identify microsatellites in mitochondrial genomes of the order Hypnales (Anand et al, 2019), *Aneura pinguis* (Kumar & Shanker, 2020a), and *Orthotrichum* (Kumar et al, 2020).

Recent advances in database development prove to be a useful resource in many scientific studies (Kumar & Shanker, 2018b, 2018c, 2020b) including characterization of microsatellites (Kumar et al, 2014; Kabra et al, 2016). In view of the immense significance of microsatellites, many specialized databases were developed including Cotton Marker Database (Blenda et al, 2006), EuMicroSatdb (Aishwarya et al, 2007), ChloroMitoSSRDB (Sablok et al, 2013), PIPEMicroDB (Sarika et al, 2013), MitoSatPlant (Kumar et al, 2014), CyanoSat (Kabra et al, 2016), PineElm_SSRdb (Chaudhary et al, 2016), and SSRome (Mokhtar & Atia, 2019).

However, available databases do not provide comprehensive information on common, polymorphic (showing length variation), and unique mitochondrial microsatellites (mtSSRs) between each pair of organisms among a genus. Therefore, we have developed a user-friendly database of pre-mined common, polymorphic, and unique mtSSRs named pSATdb (polymorphic microSATellites

[1]Department of Bioinformatics, Central University of South Bihar, Gaya, India [2]Translational Bioinformatics Lab, Department of Life Sciences, Shiv Nadar University, Greater Noida, India

Correspondence: ashutosh.bio@gmail.com; ashomics@gmail.com



pSATdb: A database of mitochondrial common, polymorphic, and unique microsatellites

HOME | TUTORIAL | STATISTICS | CONTACT

● Metazoa ○ Viridiplantae

Search genus

List of 1535 genus

Abbottina
Abscondita
Abudefduf
Acanthaspis
Acanthaster
Acanthocobitis
Acanthopagrus
Acanthorhodeus

Microsatellites, also known as simple sequence repeats (SSRs), are repetitive tracts in DNA typically consist of 1-6 nucleotides. These repeats are found in almost all organisms genome including chloroplast and mitochondria. Microsatellites are widely used molecular markers and play important role in many areas of biological research.

To analyze length variation of microsatellites between species; we have developed a database named **p**olymorphic micro**SAT**ellites **d**ata**b**ase (pSATdb). The pSATdb contains information about **Common** (found in more than one organism), **Polymorphic** (showing length variation between organisms), and **Unique** (found only in one organism) microsatellites in mitochondrial genome sequences (mtSSRs).

○ Metazoa ● Viridiplantae

Download

Search genus

List of 41 genus

Ammopiptanthus
Anomodon
Anthoceros
Bartramia
Beta
Bracteacoccus
Brassica
Calypogeia
Caulerpa
Chlamydomonas
Chlorella
Chloroparvula
Chloropicon

Show common, polymorphic, and unique mtSSRs

Frequency of mono-hexa repeats identified in 4 mitochondrial genomes of genus *Calypogeia* (Viridiplantae)

S. No.	Organisms	Accession	Repeats						Total
			Mono	Di	Tri	Tetra	Penta	Hexa	
1.	*Calypogeia arguta*	NC_035978.1	7	9	5	10	0	0	31
2.	*Calypogeia integristipula*	NC_035977.1	2	10	5	7	0	0	24
3.	*Calypogeia neogaea*	NC_035980.1	3	9	4	7	0	0	23
4.	*Calypogeia suecica*	NC_035979.1	2	9	4	8	0	0	23
Total			14	37	18	32	0	0	101

Show Flanking | Download | Back

List of 101 repeats identified in mitochondrial genomes of genus *Calypogeia* (Viridiplantae)

S. No.	Organisms	Accession	Motif	Length	Start	End
1.	*Calypogeia arguta*	NC_035978.1	(AT)$_6$	12	3796	3807
2.	*Calypogeia arguta*	NC_035978.1	(CTCC)$_3$	12	11818	11829
3.	*Calypogeia arguta*	NC_035978.1	(CAAT)$_3$	12	12229	12240
4.	*Calypogeia arguta*	NC_035978.1	(T)$_{12}$	12	14644	14655
5.	*Calypogeia arguta*	NC_035978.1	(TA)$_7$	14	16416	16429
6.	*Calypogeia arguta*	NC_035978.1	(AT)$_6$	12	16687	16698
7.	*Calypogeia arguta*	NC_035978.1	(CAAA)$_3$	12	25994	26005
8.	*Calypogeia arguta*	NC_035978.1	(TAT)$_5$	15	26791	26805
9.	*Calypogeia arguta*	NC_035978.1	(TA)$_6$	12	31504	31515
10.	*Calypogeia arguta*	NC_035978.1	(AT)$_{14}$	28	37340	37367
11.	*Calypogeia arguta*	NC_035978.1	(CCG)$_4$	12	37883	37894
12.	*Calypogeia arguta*	NC_035978.1	(AAT)$_4$	12	38673	38684

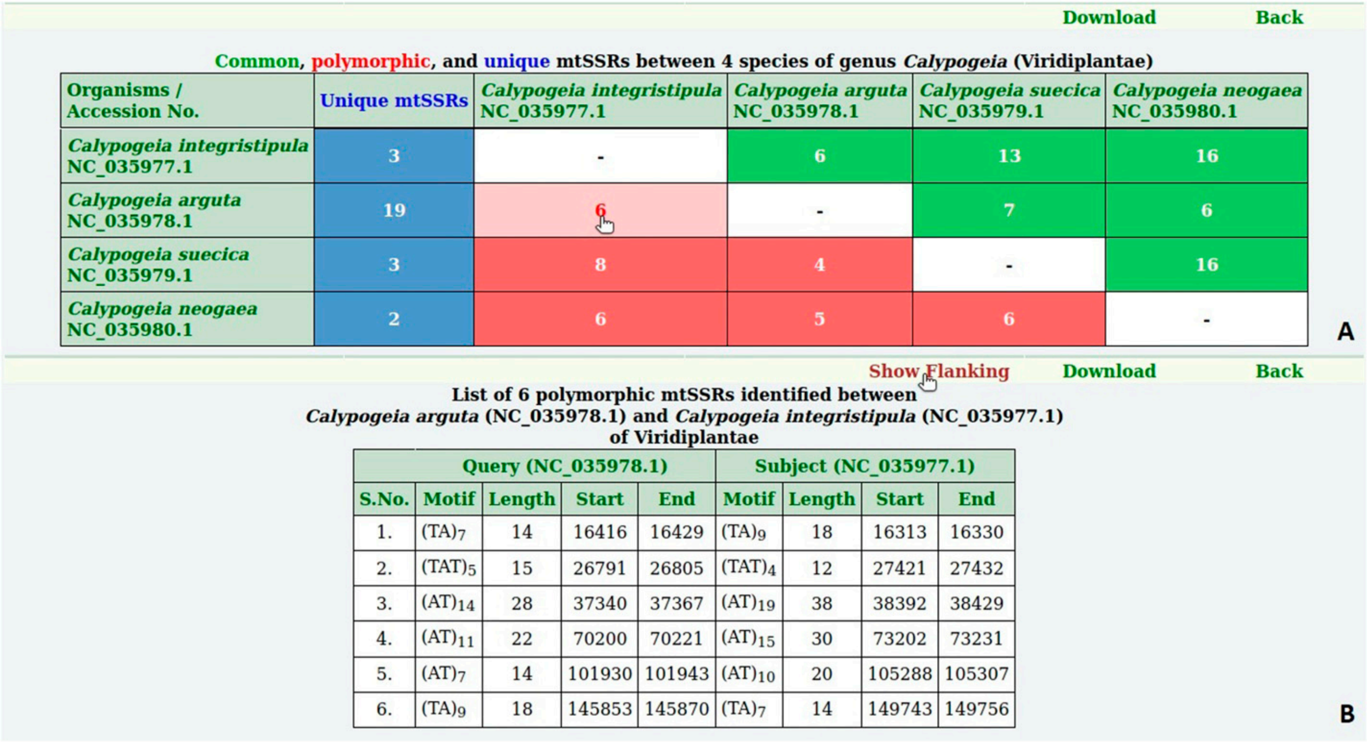
Hide Flanking | Download | Back

List of 101 repeats identified in mitochondrial genomes of genus *Calypogeia* (Viridiplantae)

Length*: Length of repeat with flanking sequences

S. No.	Organisms	Accession	Motif	Length	Start	End	Length*	Flanking Sequence
1.	*Calypogeia arguta*	NC_035978.1	(AT)$_6$	12	3796	3807	412	ACTGATAAGTTTTTGTAAGAATCTATATTTT
2.	*Calypogeia arguta*	NC_035978.1	(CTCC)$_3$	12	11818	11829	412	AAGGCTTTAGCTCTCCGCAGGCCCGTAGT
3.	*Calypogeia arguta*	NC_035978.1	(CAAT)$_3$	12	12229	12240	412	ACTCTTCTACCTAGATATACGAAAATTGGA
4.	*Calypogeia arguta*	NC_035978.1	(T)$_{12}$	12	14644	14655	412	TAAAAGCGACAGACTGTAAATCTGCTGAA
5.	*Calypogeia arguta*	NC_035978.1	(TA)$_7$	14	16416	16429	414	GAAAAAATATATATAGAATCTTCTGCCCGTA
6.	*Calypogeia arguta*	NC_035978.1	(AT)$_6$	12	16687	16698	412	AATAGCCAACTAGTGCTTGTAGCTCAATTG
7.	*Calypogeia arguta*	NC_035978.1	(CAAA)$_3$	12	25994	26005	412	TGCTTCCGCGGATGAGATGATCCGTCGAT
8.	*Calypogeia arguta*	NC_035978.1	(TAT)$_5$	15	26791	26805	415	TTGGTATATATCGCCCGCCAATCATTGGTA
9.	*Calypogeia arguta*	NC_035978.1	(TA)$_6$	12	31504	31515	412	TATATTTTGATAAACAGAGCAATCCTTTTTC
10.	*Calypogeia arguta*	NC_035978.1	(AT)$_{14}$	28	37340	37367	428	CCCGTCTTGCGTGGCTCCGGCTGTTCAGC
11.	*Calypogeia arguta*	NC_035978.1	(CCG)$_4$	12	37883	37894	412	TGAATGGAAAAAGGGCGCTTTAGATTGGG

Figure 1. The polymorphic microSATellites database and data access. (A) Home page of the polymorphic microSATellites database. **(B)** Mono-/hexa-nucleotide repeats identified in the genus *Calypogeia*. **(C)** Information on repeat motifs in forward (+ve) direction. **(D)** Repeat motifs along with their flanking sequences.



Common, polymorphic, and unique mtSSRs between 4 species of genus *Calypogeia* (Viridiplantae)

Organisms / Accession No.	Unique mtSSRs	*Calypogeia integristipula* NC_035977.1	*Calypogeia arguta* NC_035978.1	*Calypogeia suecica* NC_035979.1	*Calypogeia neogaea* NC_035980.1
Calypogeia integristipula NC_035977.1	3	-	6	13	16
Calypogeia arguta NC_035978.1	19	6	-	7	6
Calypogeia suecica NC_035979.1	3	8	4	-	16
Calypogeia neogaea NC_035980.1	2	6	5	6	-

List of 6 polymorphic mtSSRs identified between *Calypogeia arguta* (NC_035978.1) and *Calypogeia integristipula* (NC_035977.1) of Viridiplantae

	Query (NC_035978.1)				Subject (NC_035977.1)			
S.No.	Motif	Length	Start	End	Motif	Length	Start	End
1.	(TA)$_{7}$	14	16416	16429	(TA)$_{9}$	18	16313	16330
2.	(TAT)$_{5}$	15	26791	26805	(TAT)$_{4}$	12	27421	27432
3.	(AT)$_{14}$	28	37340	37367	(AT)$_{19}$	38	38392	38429
4.	(AT)$_{11}$	22	70200	70221	(AT)$_{15}$	30	73202	73231
5.	(AT)$_{7}$	14	101930	101943	(AT)$_{10}$	20	105288	105307
6.	(TA)$_{9}$	18	145853	145870	(TA)$_{7}$	14	149743	149756

Figure 2. Length variation in polymorphic microSATellites database. (A) Summary of common, polymorphic, and unique microsatellites identified in the genus *Calypogeia*. **(B)** Information on polymorphic microsatellites.

database). This database provides genus-specific information on common, polymorphic, and unique mtSSRs and can be utilized for various purposes including genetic diversity, phylogenetic analysis, and species identification.

Results

Microsatellite data access

The pSATdb (https://lms.snu.edu.in/pSATdb/) contains genus-wise information on 28,710 perfect microsatellites identified from 5,976 mitochondrial genomes of 1,576 genera which include 1,535 (5,846 mt-genome) and 41 (130 mt-genome) genera of Metazoa and Viridiplantae, respectively. Therefore, the data stored in pSATdb were categorized as Metazoa and Viridiplantae. The framework of pSATdb contains Home, Tutorial, Statistics, and Contact web pages. The "Home" page of the pSATdb provides complete access to the database. A list of genera specific to Metazoa and Viridiplantae can be retrieved by selecting the respective radio buttons. To find the desired genus, a text-based search is also provided (Fig 1A).

The frequency of repeats identified in mt-genome sequences of a genus can be fetched in a tabular form by clicking on the genus name. It will retrieve the frequency of mono-/hexa-nucleotide repeats identified in each mt-genome sequence of the selected genus (Fig 1B). Moreover, various details including repeat motif, length, and start-end position can be fetched by clicking on the respective frequency (Fig 1C). Additionally, flanking sequences of selected repeats can also be retrieved (Fig 1D).

Common, polymorphic, and unique microsatellites of the selected genus can be accessed by clicking the hyperlink "Show common, polymorphic, and unique mtSSRs" (Fig 1B). The common and polymorphic microsatellites identified between each pair of species in the selected genus were represented in the form of a matrix, whereas the total number of unique microsatellites identified in each species of selected genus was also shown in the second column of this matrix (Fig 2A).

The details of common, polymorphic, and unique microsatellites can be fetched by clicking on the respective number. It will display the repeat motif, length, and start–end position of the selected microsatellite (Fig 2B). The data stored in pSATdb can be freely downloaded using the download link.

The "Tutorial" page of pSATdb describes the functionality and interpretation of the available data. The "Statistics" page shows information about the total number of genera and species related to Metazoa and Viridiplantae available in the database. The "Contact" page is for sending any suggestion to the developers of pSATdb.

Database statistics

The pSATdb includes 1,535 genera of Metazoa and 41 genera of Viridiplantae (Fig 3A). Among all mt-genome sequences of Metazoa and Viridiplantae considered, tetranucleotides (10,323; 35.96%) were the most prevalent, followed by tri- (6,579; 22.92%), di- (4,750;

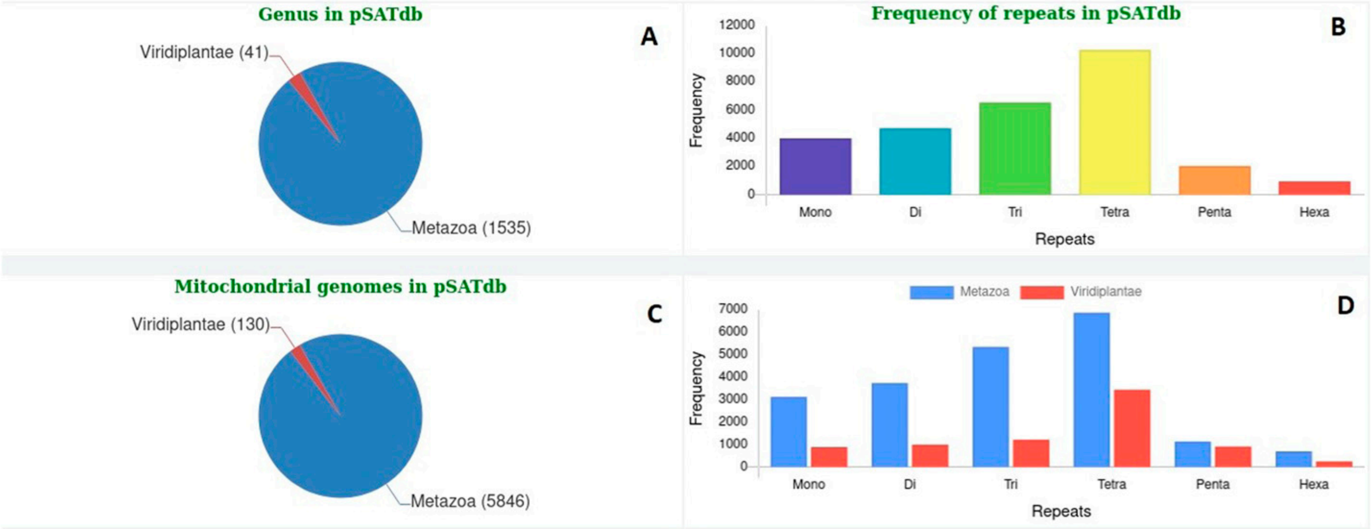


Figure 3. Details of data stored in polymorphic microSATellites database. (A) Genera of Metazoa and Viridiplantae available in polymorphic microSATellites database (pSATdb). **(B)** Frequency of mono-/hexa-nucleotide repeats in pSATdb. **(C)** Species of Metazoa and Viridiplantae included in pSATdb. **(D)** Frequency of mono-/hexa-nucleotide repeats in Metazoa and Viridiplantae.

16.54%), mono- (4,026; 14.02%), penta- (2,065; 7.19%), and hexa-nucleotide (967; 3.37%) repeats (Fig 3B).

In total, 20,960 microsatellites were identified across 5,846 mt-genomes of Metazoa (Fig 3C). Tetranucleotides (6,875; 32.80%) were the most abundant, followed by tri- (5,354; 25.54%), di- (3,745; 17.87%), mono- (3,131; 14.94%), penta- (1,146; 5.47%), and hexa-nucleotide (709; 3.38%) repeats (Fig 3D).

From 130 mt-genomes of Viridiplantae (Fig 3C), a total of 7,750 microsatellites were identified, with highest frequencies of tetra-nucleotides (3,448; 44.49%), followed by tri- (1,225; 15.81%), di- (1,005; 12.97%), penta- (919; 11.90%), mono- (895; 11.55%), and hexa-nucleotide (258; 3.33%) repeats (Fig 3D).

Common microsatellites were frequently identified between mt-genomes of closely related species (same genus). The mined data indicated that identified common, polymorphic, and unique microsatellites were not evenly distributed because of the mitochondrial genome composition and size in genera of both Metazoa and Viridiplantae (Table 1).

Table 1. Total number of genomes containing common, polymorphic, and unique microsatellites.

Genus (mt-genomes)	Common	Polymorphic	Unique
Metazoa			
536 (1,441)	—	—	✓
460 (1,907)	✓	-	✓
271 (1,627)	✓	✓	✓
100 (312)	✓	—	—
70 (177)	—	—	—
58 (209)	—	✓	✓
29 (149)	✓	✓	—
11 (24)	—	✓	—
Viridiplantae			
24 (84)	✓	✓	✓
10 (30)	—	—	✓
5 (12)	✓	—	✓
2 (4)	✓	✓	—

Discussion

In this study, microsatellites were identified in mitochondrial genomes of Metazoa and Viridiplantae and further categorized based on their genus as common, polymorphic, and unique. Earlier, mtSSRs were identified in various plants including order Hypnales (Anand et al, 2019), *Aneura pinguis* (L.) Dumort (Kumar & Shanker, 2020a), and *Orthotrichum* (Kumar et al, 2020). Apart from these, SSRs were also mined in chloroplast genomes of *Arabidopsis* (Kumar & Shanker, 2018a) and *Nymphaea* (Kumar & Shanker, 2020c). In all these studies, the distribution of mono-/hexa-nucleotide repeat motifs also varied from species to species, which is congruent with the present study. Earlier, Kumar et al (2014) observed abundance of tetranucleotide repeats in 92 organisms of Viridiplantae, and results of the present analysis are consistent with it.

Common, polymorphic, and unique mtSSRs identified in this study were not equally distributed among each genus of Metazoa and Viridiplantae. The findings are also in harmony with the length variation of microsatellites detected between each pair of species in the genus *Triticum* (Kapil et al, 2014), genus *Arabidopsis* (Kumar & Shanker, 2018a), Order Hypnales (Anand et al, 2019), and genus *Orthotrichum* (Kumar et al, 2020).

Nowadays, information on microsatellites in the public database is growing. Earlier, many databases dedicated to SSRs including

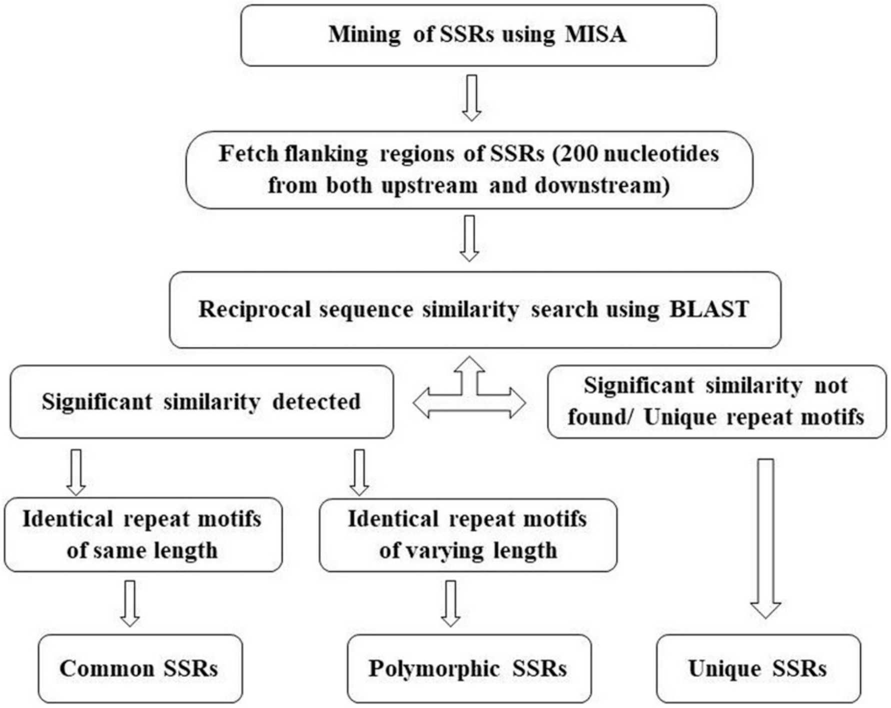

Mining of SSRs using MISA
Fetch flanking regions of SSRs (200 nucleotides from both upstream and downstream)
Reciprocal sequence similarity search using BLAST
Significant similarity detected
Significant similarity not found/ Unique repeat motifs
Identical repeat motifs of same length
Identical repeat motifs of varying length
Common SSRs
Polymorphic SSRs
Unique SSRs

Figure 4. Schematic representation to identify common, polymorphic, and unique microsatellites.

Cotton Marker Database (Blenda et al, 2006), EuMicroSatdb (Aishwarya et al, 2007), ChloroMitoSSRDB (Sablok et al, 2013), PIPE-MicroDB (Sarika et al, 2013), MitoSatPlant (Kumar et al, 2014), CyanoSat (Kabra et al, 2016), PineElm_SSRdb (Chaudhary et al, 2016), and SSRome (Mokhtar & Atia, 2019) were constructed. However, these databases lack information on common, polymorphic, and unique microsatellites. Therefore, pSATdb was developed to present information on common, polymorphic, and unique microsatellites.

Materials and Methods

Data mining

Mitochondrial genome sequences of Metazoa (animals) and Viridiplantae (plants) were downloaded from the National Center for Biotechnology Information in the FASTA file format. Initially, perfect mtSSRs were mined in retrieved mt-genomes with the help of the MIcroSAtellite Identification Tool (https://webblast.ipk-gatersleben.de/misa/; Thiel et al, 2003). The minimum repeat length of ≥12 for mononucleotide, ≥6 for dinucleotide, ≥4 for trinucleotide, and ≥3 for tetra-, penta-, and hexa-nucleotides were considered to mine the microsatellites. Moreover, interruption between two microsatellites was considered as 0.

Detection of common, polymorphic, and unique mtSSRs

Length variation between mined mtSSRs was detected using in-house–developed Perl scripts. A reciprocal similarity search was performed using the Basic Local Alignment Search Tool (Altschul et al, 1997) to establish homologous relationship between sequences containing mtSSR and, 200 base pairs of flanking sequences from both upstream and downstream of microsatellites or all nucleotides if <200 (Kabra et al, 2016; Kumar & Shanker, 2018a; Kumar et al, 2020). Microsatellites having identical repeating units with equal length and showing significant sequence similarity were categorized as common mtSSRs (found in more than one organism), whereas identical repeating units with unequal length and showing significant sequence similarity were categorized as polymorphic mtSSRs (showing length variation between organisms of a genus).

Other repeat motifs and identical repeat motifs showing no significant similarity of flanking sequences with any of the species in the same genus were considered as unique microsatellites (Kumar & Shanker, 2020a, 2020c). A schematic representation to detect common, polymorphic, and unique microsatellites is presented in Fig 4.

Database development

The pSATdb is a relational database developed using MySQL (v5.5.62). The user interface was designed in HyperText Markup Language along with Cascading Style Sheets, which were used to add style to the database. In the backend, PHP, JavaScript, and AJAX were used. Moreover, JavaScript library CanvasJS and Chart.js were used to generate the graphs.

Conclusion

A user-friendly, comprehensive database of mitochondrial microsatellites named pSATdb was successfully developed for Metazoa and Viridiplantae. It will act as a ready reference to know the length variation of repeats along with common and unique mitochondrial microsatellites within a genus. We hope that pSATdb will aid researchers working in related fields including molecular marker development, species identification, sequence-tagged sites mapping based on mitochondrial microsatellites.

Data Availability

The data available in pSATdb (https://lms.snu.edu.in/pSATdb/) are freely accessible/downloadable.

Supplementary Information

Supplementary Information is available at https://doi.org/10.26508/lsa.202101307.

Acknowledgements

S Kumar acknowledges the University Grants Commission, New Delhi, India, for financial support through the Rajiv Gandhi National Fellowship.

Author Contributions

S Kumar: resources, investigation, and writing—original draft.
A Singh: resources, methodology, and writing—review and editing.
A Shanker: conceptualization, resources, supervision, investigation, methodology, project administration, and writing—review and editing.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep* 27: 617–631. doi:10.1007/s00299-008-0507-z

Aishwarya V, Grover A, Sharma PC (2007) EuMicroSatdb: A database for microsatellites in the sequenced genomes of eukaryotes. *BMC Genomics* 8: 225–228. doi:10.1186/1471-2164-8-225

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402. doi:10.1093/nar/25.17.3389

Anand K, Kumar S, Alam A, Shankar A (2019) Mining of microsatellites in mitochondrial genomes of order Hypnales (Bryopsida). *Plant Sci Today* 6: 635–638. doi:10.14719/pst.2019.6.sp1.697

Bachmann L, Bareiss P, Tomiuk J (2004) Allelic variation, fragment length analyses and population genetic models: A case study on Drosophila microsatellites. *J Zoolog Syst Evol Res* 42: 215–223. doi:10.1111/j.1439-0469.2004.00275.x

Blenda A, Scheffler J, Scheffler B, Palmer M, Lacape JM, Yu JZ, Jesudurai C, Jung S, Muthukumar S, Yellambalase P, et al (2006) CMD: A cotton microsatellite database resource for Gossypium genomics. *BMC Genomics* 7: 132. doi:10.1186/1471-2164-7-132

Chaudhary S, Mishra BK, Vivek T, Magadum S, Yasin JK (2016) PineElm_SSRdb: A microsatellite marker database identified from genomic, chloroplast, mitochondrial and EST sequences of pineapple (Ananas comosus (L.) Merrill). *Hereditas* 153: 16. doi:10.1186/s41065-016-0019-8

Deng T, Pang C, Lu X, Zhu P, Duan A, Tan Z, Huang J, Li H, Chen M, Liang X (2016) De novo transcriptome assembly of the Chinese swamp buffalo by RNA sequencing and SSR marker discovery. *PLoS One* 11: e0147132. doi:10.1371/journal.pone.0147132

Fu D, Ma L, Qin Y, Liu M, Zhao H, Zhu G, Fu H (2016) Phylogenetic relationships among five species of Armeniaca Scop. (Rosaceae) using microsatellites (SSRs) and capillary electrophoresis. *J For Res* 27: 1077–1083. doi:10.1007/s11676-016-0245-y

Habano W, Nakamura S, Sugai T (1998) Microsatellite instability in the mitochondrial DNA of colorectal carcinomas: Evidence for mismatch repair systems in mitochondrial genome. *Oncogene* 17: 1931–1937. doi:10.1038/sj.onc.1202112

Kabra R, Kapil A, Attarwala K, Rai PK, Shanker A (2016) Identification of common, unique and polymorphic microsatellites among 73 cyanobacterial genomes. *World J Microbiol Biotechnol* 32: 71. doi:10.1007/s11274-016-2061-0

Kapil A, Rai PK, Shanker A (2014) ChloroSSRdb: A repository of perfect and imperfect chloroplastic simple sequence repeats (cpSSRs) of green plants. *Database (Oxford)* 2014: bau107. doi:10.1093/database/bau107

Kern EMA, Kim T, Park J-K (2020) The mitochondrial genome in nematode phylogenetics. *Front Ecol Evol* 8: 250. doi:10.3389/fevo.2020.00250

Kumar M, Kapil A, Shanker A (2014) MitoSatPlant: Mitochondrial microsatellites database of viridiplantae. *Mitochondrion* 19: 334–337. doi:10.1016/j.mito.2014.02.002

Kumar S, Kumari S, Shanker A (2020) In silico mining of simple sequence repeats in mitochondrial genomes of genus Orthotrichum. *J Sci Res* 64: 179–182. doi:10.37398/jsr.2020.640225

Kumar S, Shanker A (2018a) Common, unique and polymorphic simple sequence repeats in chloroplast genomes of genus Arabidopsis. *Vegetos* 31: 125–131. doi:10.5958/2229-4473.2018.00043.5

Kumar S, Shanker A (2018b) Bioinformatics resources for the stress biology of plants. In *Biotic and Abiotic Stress Tolerance in Plants*. Vats S (ed.). pp 367–386. Singapore: Springer. doi:10.1007/978-981-10-9029-5_14

Kumar S, Shanker A (2018c) Biological databases for medicinal plant research. In *Biotechnological Approaches for Medicinal and Aromatic Plants: Conservation, Genetic Improvement and Utilization*. Kumar N (ed.). pp 655–665. Singapore: Springer. doi:10.1007/978-981-13-0535-1_29

Kumar S, Shanker A (2020a) Analysis of microsatellites in mitochondrial genome of Aneura pinguis (L.) Dumort. In *Contemporary Research on Bryophytes*. Alam A (ed.), Vol. 1. pp 87–94. Singapore: Bentham Science Publishers. doi:10.2174/9789811433788120010011

Kumar S, Shanker A (2020b) Computational resources for bryology. In *Recent Advance in Botachenical Sciences: Contemporary Research on Bryophytes*. Alam A (ed.), Vol. 1. pp 20–37. Singapore: Bentham Science Publishers. doi:10.2174/9789811433788120010007

Kumar S, Shanker A (2020c) In silico comparative analysis of simple sequence repeats in chloroplast genomes of genus Nymphaea. *J Sci Res* 64: 186–192. doi:10.37398/jsr.2020.640127

Mokhtar MM, Atia MAM (2019) SSRome: An integrated database and pipelines for exploring microsatellites in all organisms. *Nucleic Acids Res* 47: D244–D252. doi:10.1093/nar/gky998

Parida SK, Kalia SK, Kaul S, Dalal V, Hemaprabha G, Selvi A, Pandit A, Singh A, Gaikwad K, Sharma TR, et al (2009) Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theor Appl Genet* 118: 327–338. doi:10.1007/s00122-008-0902-4

Powell W, Machray G, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1: 215–222. doi:10.1016/s1360-1385(96)86898-0

Roger AJ, Muñoz-Gómez SA, Kamikawa R (2017) The origin and diversification of mitochondria. *Curr Biol* 27: R1177–R1192. doi:10.1016/j.cub.2017.09.015

Sablok G, Mudunuri SB, Patnana S, Popova M, Fares MA, Porta NL (2013) ChloroMitoSSRDB: Open source repository of perfect and imperfect repeats in organelle genomes for evolutionary genomics. *DNA Res* 20: 127–133. doi:10.1093/dnares/dss038

Sarika S, Arora V, Iquebal MA, Rai A, Kumar D (2013) PIPEMicroDB: Microsatellite database and primer generation tool for pigeonpea genome. *Database (Oxford)* 2013: bas054. doi:10.1093/database/bas054

Shanker A, Bhargava A, Bajpai R, Singh S, Srivastava S, Sharma V (2007a) Bioinformatically mined simple sequence repeats in UniGene of Citrus sinensis. *Sci Hortic* 113: 353–361. doi:10.1016/j.scienta.2007.04.011

Shanker A, Singh A, Sharma V (2007b) In silico mining in expressed sequences of Neurospora crassa for identification and abundance of microsatellites. *Microbiol Res* 162: 250–256. doi:10.1016/j.micres.2006.05.012

Stoeckle MY, Thaler DS (2018) Why should mitochondria define species? *Hum Evol* 33: 1–30. doi:10.1101/276717

Stolle E, Kidner JH, Moritz RF (2013) Patterns of evolutionary conservation of microsatellites (SSRs) suggest a faster rate of genome evolution in Hymenoptera than in Diptera. *Genome Biol Evol* 5: 151–162. doi:10.1093/gbe/evs133

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res* 12: 4127–4138. doi:10.1093/nar/12.10.4127

Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). *Theor Appl Genet* 106: 411–422. doi:10.1007/s00122-002-1031-0

Yang L, Tan Z, Wang D, Xue L, Guan MX, Huang T, Li R (2014) Species identification through mitochondrial rRNA genetic analysis. *Sci Rep* 4: 4089. doi:10.1038/srep04089